


[Web](#) [Images](#) [Video](#) [News](#) [Maps](#) [more »](#)


[Advanced Scholar Search](#)  
[Scholar Preferences](#)  
[Scholar Help](#)

## Scholar All articles - Recent articles

 Results 1 - 10 of about 41,300 for **web crawlers**. (0.23 seconds)

**Mercator: A scalable, extensible web crawler-** ► [uwo.ca](#) [PDF]

A Heydon, M Najork - *World Wide Web*, 1999 - Springer

... Scalable **Web crawlers** are an important component of many **Web services**, but their design is not well-documented in the literature. ...

Cited by 263 - Related articles - Web Search - BL Direct - All 49 versions

**Evaluating topic-driven Web crawlers-** ► [case.edu](#) [PDF]

F Menczer, G Pant, P Srinivasan, ME Ruiz - *Proceedings of the 24th annual international ACM SIGIR ...*, 2001 - portal.acm.org

Page 1. Evaluating Topic-Driven **Web Crawlers** ... Although **crawlers** related to general-purpose **Web search engines** are also important, we are not focusing on these. ...

Cited by 146 - Related articles - Web Search - All 21 versions

**SPHINX: a framework for creating personal, site-specific web crawlers-** ► [toronto.edu](#) [PS]

RC Miller, K Bharat - *Computer Networks and ISDN systems*, 1998 - Elsevier

... SPHINX: a framework for creating personal, site-specific **Web crawlers**. ... Stateofhearth

**Web crawlers** are generally handcoded programs in Perl, C C++, or Java. ...

Cited by 140 - Related articles - Web Search - All 5 versions

[PDF] ► **Ubicrawler: A scalable fully distributed web crawler**

P Boldi, B Codenotti, M Santini, S Vigna - *Software: Practice and Experience*, 2004 - gopinaras.50webs.com

... downloading pages in a reasonable amount of time\* [9, 1]. Many commercial and research institutions run their **web crawlers** to gather data about the **web**. ...

Cited by 119 - Related articles - View as HTML - Web Search - BL Direct - All 29 versions

**A Web crawler design for data mining-** ► [psu.edu](#) [PDF]

M Thelwall - *Journal of Information Science*, 2001 - jis.sagepub.com

... For this reason, **crawlers** are normally multi-threaded, so that hundreds of **web pages** may be requested simultaneously by one process [23, 24]. ...

Cited by 77 - Related articles - Web Search - All 6 versions

**Topical web crawlers: Evaluating adaptive algorithms-** ► [nyu.edu](#) [PDF]

F Menczer, G Pant, P Srinivasan - *ACM Transactions on Internet Technology (TOIT)*, 2004 - portal.acm.org

Page 1. Topical **Web Crawlers: Evaluating Adaptive Algorithms** ... Page 2. Topical

**Web Crawlers: Evaluating Adaptive Algorithms** - 379 ...

Cited by 93 - Related articles - Web Search - All 14 versions

**Effective page refresh policies for web crawlers-** ► [stanford.edu](#) [PDF]

J Cho, H Garcia-Molina - *ACM Transactions on Database Systems (TODS)*, 2003 - portal.acm.org

Page 1. Effective Page Refresh Policies for **Web Crawlers** ... Page 2. Effective Page Refresh Policies for **Web Crawlers** - 391 users access relevant information. ...

Cited by 75 - Related articles - Web Search - BL Direct - All 16 versions

**Finding near-replicas of documents on the web-** ► [psu.edu](#) [PDF]

N Shivakumar, H Garcia-Molina - *Lecture notes in computer science*, 1999 - Springer

... This information can be used to improve **web crawlers**, **web archivers** and in the presentation of search results, among others. We ...

Cited by 131 - Related articles - Web Search - BL Direct - All 20 versions

**MySpiders: Evolve your own intelligent Web crawlers-** ► [psu.edu](#) [PDF]

G Pant, F Menczer - *Autonomous agents and multi-agent systems*, 1999 - Springer

MySpiders: Evolve Your Own Intelligent **Web Crawlers** GAUTAM PANT AND FILIPPO MENCZER ...

"Evaluating topic-driven **web crawlers**," in *Proc. 24th Annual Int. ...*

Cited by 47 - Related articles - Web Search - BL Direct - All 6 versions

An adaptive model for optimizing performance of an incremental web crawler- ► [psu.edu \[PDF\]](#)

J Edwards, K McCurley, J Tomlin - ... of the 10th international conference on World Wide Web, 2001 - [portal.acm.org](#)

Page 1. An Adaptive Model for Optimizing Performance of an Incremental Web

Crawler Jenny Edwards £ Faculty of Information Technology ...

Cited by 112 - Related articles - Web Search - All 19 versions

Key authors: [H Garcia-Molin...](#) - [M Najork](#) - [J Cho](#) - [A Heydon](#) - [F Menczer](#)

Google

Result Page: 1 2 3 4 5 6 7 8 9 10 **Next**

[Google Home](#) - [About Google](#) - [About Google Scholar](#)

©2009 Google